UNITED STATES PATENT APPLICATION


OF


**Adrian Evans CONWAY**


**FOR**


**A PASSIVE SYSTEM AND METHOD FOR MEASURING THE SUBJECTIVE QUALITY OF REAL-TIME MEDIA STREAMS IN A PACKET-SWITCHING NETWORK**

## BACKGROUND OF THE INVENTION

### Field of the Invention

[0001] A system and method are disclosed for the non-intrusive determination of the
subjective quality of a packetized real-time media data stream in a packet-switching
network without interrupting the flow of the media stream or adding a known test signal
in the packet-switching network.

### Description of the Related Art

[0002] In recent years, packet-switching networks have been used increasingly for
transport of real-time media such as, for example, voice and video media signals that are
transmitted either real-time or delayed. The signals begin as either analog or digital
signals. A specific example is the increasing use of the Internet for carrying voice-over-
internet protocol (VoIP) calls. In a VoIP call, a digitally encoded voice signal is
packetized into an Internet Protocol (IP) packet stream that is transmitted over the
Internet to a destination device. At the destination device, the digitally encoded voice
signal is extracted from VoIP packet payloads in the packet stream and then decoded into
a signal that is played-out in real-time to the user at the destination.

[0003] When a packetized real-time media stream is transmitted across a packet-
switching network, the packet stream may be corrupted by a number of network
impairments. Examples of network impairments include packet-discarding at routers due
to packet bit errors, packet-dropping at interface buffers due to traffic congestion, packet-
duplication, packet-delay in time beyond its hard or soft real-time deadline at its
destination, packet-misrouting, and loss of packet-sequence. These impairments generally
degrade the quality of the media signal that is eventually received at the destination.

[0004] Due to network impairments that may be encountered in the transmission of
real-time media over a packet-switching network, it is important to be able to measure
and monitor the quality-of-service (QoS) that is being provided by the network. Typical
network QoS measures include, for example, end-to-end packet delay, end-to-end packet

delay jitter, packet corruption, and packet loss.  To monitor such network QoS measures, one can deploy commercially available monitoring systems.

[0005]    Although measurement and monitoring of network QoS provides valuable information regarding the ability of a network to properly support transmission of real-time media signals, such measures do not directly reflect the media signal subjective quality that is actually perceived by an end-user.  This is the case since subjective quality of a real-time media signal, as perceived by the end-user, is difficult to quantify in terms of the network QoS measures.  To deal with this general problem, objective methods have been developed for estimating subjective quality of media signals.  For example, perceptual speech quality measurement, PSQM, is a means for objectively assessing the quality of speech that has been degraded by a telephony network.  It has a high correlation to subjective quality across a range of distortion types, and is used to test networks that are subject to different coding types and transmission errors.  PSQM is used primarily to test networks that have speech compression, digital speech interpolation, and packetization.  PSQM of this type has been recommended by the International Telecommunication Union in its telephone recommendations.  An example of an objective method for voice signals is ITU-T recommendation P.861.

[0006]    To estimate the subjective quality of a VoIP call that traverses a packet-switching network, Fig. 1 depicts the basic approach of PSQM that has been adopted.  In an objective method a known artificial voice signal 10 is transmitted across the network 12. An artificial voice signal for use in PSQM may be stored in a commonly used file format such as, for example, a wave (.WAV) file.  The process of transmitting a known signal to evaluate the degradation in quality after it has traversed a network may be termed an active method since a known test signal is actually injected into and transmitted across the network.

[0007]    PSQM uses a psychoacoustic model 14 that aims to mimic the perception of sound in real life, and was originally developed to test compressor/decompressors (codecs).  A Codec is a software component that translates video or audio between its

uncompressed form and the compressed form in which it is stored. The algorithm

functions by comparing signal 10 after it has been through a coder 16 and decoder 18.

[0008]    PSQM provides an output 20 in the range of 0 to 6.5, where values close to 0

indicate very good speech quality, and values close to 6.5 indicate poor speech quality. At

5    the destination, a quality measure or score (e.g. mean opinion score (MOS)) is computed

22 based on the received signal and the known artificial voice signal that was transmitted

and is output 24. Although PSQM does not have a direct correlation to MOS, the

subjective quality is nevertheless inferred from the objective quality. That is, if a person

listens to a speech sample that has a PSQM value of 2, that person would think the quality

10    was worse than a speech sample having a PSQM value of 1. PSQM values can be

roughly translated in MOS values.

[0009]    Accordingly, there is a need for an accurate subjective measurement algorithm

for determining the mean opinion score and quality of VoIP packets.


15                          <u>**SUMMARY OF THE INVENTION**</u>

[0010]    This invention provides a system and method for the non-intrusive

determination of the subjective quality of a packetized real-time media data stream in a

packet-switching network. It also provides for the subjective determination of the media

data stream without interrupting the flow of the media stream. Moreover, the invention

20    allows for the subjective quality determination of the media stream without the addition

of a known test signal into the packet-switching network. Passive subjective quality

determination of the media stream using this method prevent perturbing the packets in the

media stream, thereby increasing the strength of the subjective results.

[0011]    A media stream to be measured is transported using packets containing a

25    timestamp, a sequence number, and a source identifier. The type of encoding and

packetization method may be known from knowledge that one may have about the media

stream being measured or it may be known explicitly from information contained in the

media packet headers. For example, in the case of real-time transfer protocol (RTP)

packets, the header typically contains a 7-bit payload type field that identifies the format

of the RTP payload and determines how it should be interpreted.

[0012]    First, packets are copied from a network with their associated timestamps and

stored in a memory without interfering with the flow of data in the network. The

5    timestamps of the packets are the local time at which the packet is copied, and ensures

that temporal correlation can be performed between the original packets and the copied

packets. Copied packets are separated into individual streams of packets utilizing the

source identifier field in each packet header. The media signal in an individual copied

packetized stream is replaced with a known test signal forming a pseudo-media stream.

10    The same methods are used to encode, packetize, depacketize and decode the known test

signal as those used on the copied media stream signal. The pseudo-media stream

therefore contains all of the faults of the copied signal including any faults incurred in the

copied signal during its transmission. The subjective quality of the pseudo-media stream

is then measured disclosing the subjective quality of the data signal. These and other

15    advantages of this invention are described in or are apparent from the following detailed

description of the preferred embodiments.

<p align="center">**BRIEF DESCRIPTION OF THE DRAWINGS**</p>

[0013]    In the drawings:

[0014]    Figure 1 depicts the prior art PSQM algorithm for measuring the objective

20    quality of a media signal;

[0015]    Figure 2 shows the components of the passive system used to measure the

subjective quality of a media signal;

[0016]    Figure 3 depicts a flow chart in accordance with the preferred embodiment of

the invention;

25    [0017]    Figure 4 depicts packets of a know signal having known packet payloads;

[0018]    Figure 5 depicts packets copied from a media stream of interest; and

[0019]    Figure 6 depicts a pseudo packet media stream with the media packet payloads

of interest replaced with known media payloads.

## DESCRIPTION OF THE PREFERRED EMBODIMENT

[0020]     Figure 2 depicts a passive system 30 for measuring subjective quality of real-time media streams 32 in a packet-switching network 34.  The real-time media streams 32 are obtained from a source 31 and can be in an analog or digital.  The real-time media streams 32 are digitally encoded and packetized to form a packet stream that is transmitted over the packet-switching network.  A packet stream 36 is copied from network 34 without interfering with the data stream carried by the network.  For example, a packet stream 36 may be copied from any suitable port, not illustrated, for example, an Ethernet hub, a tap on a transmission link, or a packet-switch that is configured as a switched port analyzer port ('SPAN port').  A SPAN is a switch that extends the monitoring abilities of existing network analyzers into a switched Ethernet environment.  SPAN mirrors the traffic at one switched segment onto a predefined SPAN port.  A network analyzer 38 attached to the SPAN port can monitor traffic from any other switched ports.

[0021]     The copied portion of the packet stream is sent to a temporary storage 40 and marking unit 42 where each copied packet is given a timestamp corresponding to the current time at which it is copied.

[0022]     Memory 40 is preferably implemented using static or dynamic random access memory (RAM).  However, memory 40 can also be implemented using one or more of static RAM, dynamic RAM, read-only memory (ROM), flash memory, hard disk drive, compact disc read-only memory (CD-ROM) drive, compact disc-readable/writable (CD-RW) drive, Floppy Disk Drive, Network Servers or the like.

[0023]     Copied packets 36 are retrieved by a simulator 44, where the packets 36 are separated from the multisource media stream according to information contained in each packet header portion.  The encoding/decoding and packetization/depacketization methods used and to be used with the selected single source media stream are also determined by the simulator.

[0024]     A known media signal 46 is sent to the simulator 44 where the known media signal is encoded and packetized using the same encoding and packetization as the real-

time media streams 32, and is substituted into packets 36, and output as pseudo-packets, or pseudo-media stream 48 to a quality measurement comparator 50.

[0025]    Pseudo-packets 48 contain the known media signal 46 along with all the faults in the selected packets 36. Comparator 50 determines the subjective quality of the pseudo-packets 48, and outputs a quality measurement 52.

[0026]    Fig. 3 depicts a flow chart of the steps to determine the quality of the media signal 32 arriving from source 31, starting at step S100, and continuing to step S110.

[0027]    In step S110, media stream 36 is received from packet-switching network 34, and continues to step S120. In step S120, packets 26 are copied from network 34 without interfering with the data stream carried by the network, and are timestamped. The process continues to step S130. The copied and timestamped packets are then stored in memory 40 at step S130.

[0028]    Next, in step S140, the stored packets are separated by their media source, which are identified from source information contained in each packet header portion, and continues to step S150. In step S150, the encoding and decoding methods used by the packets 36 is determined, and the process continues to step S160.

[0029]    The method of packetization and depacketization of the packets 36 are determined in step S160. The method of encoding/packetization is output to step S170, and the method of decoding/depacketization is output to step S200, and the process continues to step S170.

[0030]    A known media signal 46 is encoded and packetized in step S170 using the same methods used in encoding and packetizing the selected media stream.

[0031]    In step S180, the encoded and packetized known signal 46 proceeds to the simulator 44 where the payloads are emptied from each copied packet and a new payload is added corresponding to the encoded and packetized known media signal. It should be noted that each loaded packet in the pseudo-media stream contains the same packet payload that would have been contained in the media packets had the media stream actually carried the known media signal encoded and packetized. The media stream containing the known media signal has all of the same faults as the copied media signals.

For example, if packets from the copied medial signal are missing for any reason or are

out of sequence, the pseudo-media stream will have the same faults.

**[0032]**      Next, in step S180, arrival of the pseudo-media packet streams is simulated,

and continues to step S200. In step S200, the pseudo-media packet streams are decoded

5     and depacketized, and continues to step S210.

**[0033]**      In step S210, the objective quality of the decoded and depacketized pseudo-

media stream is determined, using, for example, PSQM, and continues to step S220. In

step S220, the subjective quality of the decoded and depacketized pseudo-media stream is

determined, for example, by comparing the decoded and depacketized pseudo-media

10    stream to the known signal before it was inserted into the packet, and determining a

quality measurement score, and the process ends at step S230.

**[0034]**      The substitution of a known signal into a packet is further illustrated in Figs.

4-6, where a known media signal 46 is represented. The packets are, for example, real-

time transport protocol (RTP) packets, defined in Request for Comments (RFC) 1889,

15    that are commonly used to carry VoIP and other media. In an RTP packet, the header

typically includes a 32-bit timestamp, a 16-bit sequence number, and a 32-bit

Synchronization Source (SSRC) identifier. Known media signal 46 has been packetized

in representation 54 using the same method as that used to packetize the media signal 36

under measurement. In Fig. 5 it can be seen that each packet of the copied signal 36a,

20    36b, 36c, 36d, 36e and 36f is timestamped and identified by its sequence number ("SN")

in the stream. Any corrupted packets are eliminated as they would be at the receiving

end in the packet-switching network 34. It should also be noted that there are gaps in the

sequence numbers of the copied media stream. As previously mentioned, these gaps

could result from many possible problems.

25    **[0035]**      Each packet 36 has a payload portion 56 and a header portion 58 as shown in

the legend of Fig. 6. The header portion 58 includes a timestamp, a sequence number and

a source identifier. When replacing the media packet payloads the sequence numbers are

used to determine which known media signal packet payloads should replace the copied

media stream packet payloads. The packet payloads of the media stream 36 contain the

encoded digital signal. The payloads 56 of each of the copied and timestamped packets 36a-36f are then emptied and the packets of the known signal 54 are added to the packets 36a-36f to create a pseudo-media stream 48 depicted in Fig. 6.

**[0036]** In Fig. 6, the timestamped pseudo-media stream 48 is shown complete with all
5   the faults present in the copied media signal 36. Note that the packets identified by sequence number (SN) 25, 28 and 29 are missing. The pseudo-media stream 48 containing the packetized known test signal is an accurate representation of the copied media stream. Timestamping of each packet 36 upon copying assures that the signal to be tested has the same temporal sequence of packet elements as the original source media
10  signal 32.

**[0037]** The above described passive system possesses a number of inherent advantages relative to the active system of measuring the subjective quality of media streams, including the following:

First, the passive method can be used anywhere on the network to measure
15      the subjective quality of live real-time media streams that are in the process of being transmitted across the network. The active method can only be used to measure the subjective quality of test media streams that are injected into the network.

Second, the passive method can be used to measure simultaneously the
20      subjective quality of multiple live real-time media streams since the copied packet stream may contain multiple media streams that can be distinguished by their source.

Third, the passive method does not require any sending device. The active method requires a sending device for inserting a packetized test signal into the
25      network.

Fourth, the passive method may be used to measure simultaneously, in one test, the subjective quality of a media stream using more than one subjective quality measurement technique. This is because the passive method does not

require the transmission of a test signal and is not restricted to any specific subjective quality measurement technique.

Fifth, the passive method may be used to measure simultaneously the subjective quality of a media stream using more than one known media signal. This is because the passive method does not require the transmission of a test signal and is not restricted to any specific known media signal 46. In the active method, different known media signals would have to be transmitted separately to obtain subjective quality measurements for different known media signals.

Sixth, the passive method may be used to measure simultaneously the subjective quality of a media stream using more than one subjective quality measurement technique and more than one known media signal. This is because it is possible to simultaneously use more than one subjective quality measurement technique and, for each such technique, utilize more than one known media signal.

Seventh, the passive method is not restricted to any particular media type (e.g. voice, video, audio, or combinations thereof (multimedia), since no assumptions are made regarding the form of the media signal.

Eighth, the passive method may be repeated periodically in time under automatic control to enable the continuous monitoring of the subjective quality of media streams.

[0038] The above described advantages of the passive system are by no means meant to limit the scope of the invention. Though the invention has been described with respect to a single preferred embodiment thereof, many variations and modifications will become apparent to those skilled in the art. It is therefore the intention that the claims be interpreted as broadly as possible in view of the prior art, to include all such variations and modifications.